

DOCUMENT RESUME

ED 455 288

TM 033 141

AUTHOR Webb, Noreen M.; Welner, Mari; Zuniga, Stephen
TITLE Short Circuits or Superconductors? Examining Factors That Encourage or Undermine Group Learning and Collaboration among High-Ability Students. CSE Technical Report.
INSTITUTION Center for Research on Evaluation, Standards, and Student Testing, Los Angeles, CA.; California Univ., Los Angeles. Center for the Study of Evaluation.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
REPORT NO CSE-TR-541
PUB DATE 2001-05-00
NOTE 48p.
CONTRACT R305B60002
AVAILABLE FROM National Center for Research on Evaluation, CRESST/CSE, Graduate School of Education & Information Studies, University of California, Los Angeles, CA 90095-1522. Tel: 310-206-1532.
PUB TYPE Reports - Research (143)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Ability; *Group Instruction; *Heterogeneous Grouping; High Achievement; *Homogeneous Grouping; *Junior High School Students; Junior High Schools; Performance Based Assessment; Performance Factors

ABSTRACT

This study investigated the effects of group ability composition (homogeneous versus heterogeneous) on group processes and outcomes for high ability students completing science performance assessments. Participants were 99 seventh and eighth graders from 9 classes in 2 schools. The results show that group ability composition does not have straightforward effects on achievement. While high-ability students working in homogeneous groups uniformly performed well, high-ability students in some heterogeneous groups performed better than high-ability students in other heterogeneous groups. The quality of group functioning served as the strongest predictor of high-ability students' achievement. High-ability students in groups that were responsive to group members' needs for help and did not engage in debilitating social behavior performed well, whereas high-ability students in groups with poorer functioning did not. While homogeneous groups consistently showed beneficial group functioning, only some heterogeneous groups exhibited these traits. These results show that achievement of high-ability students cannot be predicted from a simple homogeneous-heterogeneous grouping contrast and that the level of group functioning provides the key to understanding group performance. An appendix contains sample test questions. (Contains 10 tables, 10 figures, and 63 references.) (SLD)

CRESST

National Center for Research on Evaluation, Standards, and Student Testing

ED 455 288

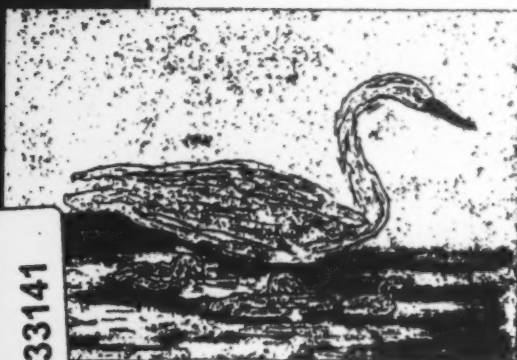
U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- This document has been reproduced as received from the person or organization originating it.
- Minor changes have been made to improve reproduction quality.
- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Short Circuits or Superconductors? Examining Factors That Encourage or Undermine Group Learning and Collaboration Among High-Ability Students

CSE Technical Report 541

Noreen M. Webb, Mari Welner, and Stephen Zuniga
CRESST/University of California, Los Angeles



TM033141



UCLA Center for the Study of Evaluation

BEST COPY AVAILABLE

**Short Circuits or Superconductors?
Examining Factors That Encourage or Undermine
Group Learning and Collaboration
Among High-Ability Students**

CSE Technical Report 541

Noreen M. Webb, Mari Welner, and Stephen Zuniga
CRESST/University of California, Los Angeles

May 2001

Center for the Study of Evaluation
National Center for Research on Evaluation,
Standards, and Student Testing
Graduate School of Education & Information Studies
University of California, Los Angeles
Los Angeles, CA 90095-1522
(310) 206-1532

Project 1.1 Models-Based Assessment Design: Individual and Group Problem Solving—Science,
Noreen Webb, Project Director, CRESST/UCLA

Copyright © 2001 The Regents of the University of California

The work reported herein was supported under the Educational Research and Development Centers
Program, PR/Award Number R305B60002, as administered by the Office of Educational Research
and Improvement, U.S. Department of Education.

The findings and opinions expressed in this report do not reflect the positions or policies of the
National Institute on Student Achievement, Curriculum, and Assessment, the Office of Educational
Research and Improvement, or the U.S. Department of Education.

**SHORT CIRCUITS OR SUPERCONDUCTORS?
EXAMINING FACTORS THAT ENCOURAGE OR UNDERMINE
GROUP LEARNING AND COLLABORATION
AMONG HIGH-ABILITY STUDENTS**

**Noreen Webb, Kariane Mari Welner, and Stephen Zuniga
CRESST/University of California, Los Angeles**

Abstract

Peer-directed small-group collaboration is featured prominently in debates about good classroom instruction and in the promotion of school reform. Although many cooperative learning methods advocate grouping students heterogeneously in order to maximize the diversity of perspectives, skills, and backgrounds, past research shows that while heterogeneous grouping generally benefits low-ability students, it does not necessarily benefit high-ability students. This study investigates the effects of group ability composition (homogeneous vs. heterogeneous) on group processes and outcomes for high-ability students completing science performance assessments. The results show that group ability composition does not have straightforward effects on achievement. While high-ability students working in homogeneous groups uniformly performed well, high-ability students in some heterogeneous groups performed better than high-ability students in other heterogeneous groups. The quality of group functioning served as the strongest predictor of high-ability students' achievement. High-ability students in groups that were responsive to group members' need for help and did not engage in debilitating social behavior performed well, whereas high-ability students in groups with poorer functioning did not. Whereas homogeneous groups consistently showed beneficial group functioning, only some heterogeneous groups exhibited these traits. These results show that achievement of high-ability students cannot be predicted from a simple homogeneous-heterogeneous grouping contrast and that the level of group functioning provides the key to understanding group performance.

Debates about good classroom instruction and the promotion of school reform prominently feature peer-directed small-group collaboration. Group collaboration can increase learning and achievement, social skills, self-esteem, and attitudes toward classmates and school. Past research demonstrates that lower ability students usually benefit from collaborative learning, at least when working in groups heterogeneous in ability or achievement. In such groups, lower ability students can learn from their more able peers. The benefits of group collaboration

for high-ability students, however, are less clear. Previous research on the benefits of collaborative work for high-ability students and the relative advantages of different group compositions for these students yields inconsistent results. Moreover, current research tells us little about the processes operating in collaborative groups that may work to the benefit or detriment of high-ability students' learning. The present study, then, examined high-ability 8th-grade science students' performance during and after solving electric circuit building problems individually and/or in collaborative groups. This study specifically investigated (a) group processes that positively or negatively influenced the achievement of high-ability students, and (b) the influence of group composition on group processes and achievement outcomes.

Group Processes Hypothesized to Promote Learning

Piaget's model of socio-cognitive conflict and learning provides a useful theoretical perspective regarding the mechanisms by which learning occurs within group contexts. Interaction with others may produce discrepancies between a child's views and new information, giving rise to cognitive conflict within the child, and leading the child to try out new ideas. Internal cognitive conflict may manifest itself in interaction with others through overt conflict or controversy in which individuals recognize the differences between their beliefs. Overt conflict encourages individuals to explain and justify their own positions, raises uncertainties about their beliefs, encourages individuals to seek new information to help resolve their disagreements and arguments, and helps them understand alternative points of view (Brown & Palincsar, 1989; Johnson & Johnson, 1979).

The co-construction of knowledge that some or all students did not have prior to collaboration provides a second way students can benefit from group work (Damon & Phelps, 1989). When a group has a student with high subject-matter knowledge, that student can help less knowledgeable peers carry out tasks or solve problems that they could not perform without assistance, a process called scaffolding (Wood, Bruner, & Ross, 1976). During this process, the learners practice skills that they then internalize, which increase their understanding and enable them to accomplish more individually (Tudge & Rogoff, 1989; Vygotsky, 1981). Even if a group contains no expert, students can build on each other's knowledge and understanding to solve problems and develop new knowledge that none could accomplish alone (Forman & Cazden, 1985; Saxe, 1992).

Students working collaboratively can also learn by helping each other. From a theoretical perspective, both the help-giver and the help-receiver stand to benefit from elaborated help, that is, detailed explanations about how to solve problems or complete tasks. Giving explanations encourages the explainer to clarify and reorganize the material in new ways in order to make it understandable to others (Bargh & Schul, 1980). This cognitive restructuring may help the explainer to understand the material better, develop new perspectives, and recognize and fill in gaps in her understanding. By tailoring explanations to the difficulties of other students, helpers may construct more elaborate conceptualizations than they would when solving the problems for themselves (Benware & Deci, 1984; Yackel, Cobb, & Wood, 1991). Receiving explanations can benefit the receiver by filling in gaps in her understanding, correcting misconceptions, and strengthening connections between new information and previous learning (Mayer, 1984; Wittrock, 1990). Students have the potential to explain effectively because peers share a similar language and can translate difficult vocabulary and expressions into words that fellow students can understand (Noddings, 1985). Further, because explainers and other students often learn material at the same time, this enables explainers to tune into other students' misconceptions and give relevant explanations (Vedder, 1985). Students can also control the pace of explanations (as well as the pace of group work) to better promote understanding. Students experience the benefits of receiving explanations most strongly when receivers have, and use, opportunities to apply the explanations to solve problems or perform tasks for themselves (Vedder, 1985; Mastergeorge, Webb, Roc, & Baure, 2000; Webb, Troper, & Fall, 1995). Carrying out further activity after receiving explanations may (a) help learners internalize principles and construct-specific inference rules for solving problems (Chi & Bassock, 1989; Chi, Bassock, Lewis, Reimann, & Glaser, 1989), (b) help learners monitor their own understanding (Chi & Bassock, 1989), and (c) help make the group aware of a student's continued misunderstandings (Shavelson, Webb, Stasz, & McArthur, 1988).

The beneficial social-emotional processes that take place in collaborative groups can also enhance learning, especially if students receive incentives to cooperate, such as rewards for high group performance or effort. According to the motivational theory of Deutsch (1949), when groups work toward a common goal, students will praise, encourage, and support each other's efforts, resulting in greater

effort and an increased liking of the task and other students (see also Johnson & Johnson, 1985).

However, not all groups function in ways that promote learning. A variety of debilitating processes may inhibit group functioning. For example, a number of studies found that receiving non-elaborated help (e.g., the answer without any description of how to obtain it) in response to questions or errors negatively affects achievement (see Webb & Palincsar, 1996). Another debilitating process concerns the failure of students to seek help when they need it. Insights into when students do and do not seek help come from Nelson-Le Gall's (1981, 1985, 1992; Nelson-Le Gall, Gumerman, & Scott-Jones, 1983) comprehensive, five-step model of children's help-seeking. For students to benefit from group work, they must realize that they need help, be willing to seek help, identify someone who can provide help, use effective strategies to elicit help (e.g., ask explicit, precise, and direct questions; Peterson, Wilkinson, Spinelli, & Swing, 1984; Wilkinson, 1985; Wilkinson & Calculator, 1982a, 1982b; Wilkinson & Spinelli, 1983; see also Webb et al., 1995; Webb & Kenderski, 1984), and be willing to reassess their strategies for obtaining help.

Debilitating socio-emotional processes can also occur within groups, decreasing the chance that students will benefit from the group-work experience. Salomon and Globerson (1989) observed four such processes among pairs of students working on reading and writing tasks at the computer. In the "free rider" effect (see also Kerr & Bruun, 1983), also called "diffusion of responsibility" (Slavin, 1990), some group members sat back and let others do the work. This sometimes produced a "sucker effect" in which the group members doing all of the work discover that their fellow students took them for a free ride and therefore start to contribute less to group work to avoid the label of "sucker." The dominance of some group members over others constituted a third effect, with high-status students having undue influence over group functioning, thereby preventing low-status students from making contributions or obtaining the help that they needed (see Dembo & McAuliffe, 1987). Finally, Salomon and Globerson observed some pairs deciding to go through the motions without actually performing the assigned task, feigning busyness. Deering (1989) observed additional disabling processes, such as aggressiveness and hostility leading to unconstructive and bitter arguments, and passivity and acquiescence leading to premature agreement on answers.

Nearly all research investigating the relationships between group processes and learning outcomes combines all ranges of students' ability. Therefore, this body

of work tells us little about the processes that specifically predict high-ability students' learning outcomes. Yet, the small amount of research that focuses specifically on high-ability students demonstrates similar patterns in what makes for a positive group experience. Swing and Peterson (1982) found that on two out of three outcome measures, giving conceptual or sequencing explanations related positively, and carrying out off-task behavior related negatively, to high-ability students' achievement.

Achievement of High-Ability Students in Homogeneous and Heterogeneous Groups

Teachers often ask how they should assign students to collaborating groups within their classrooms. Many cooperative learning methods recommend that teachers form groups heterogeneously to reflect the diversity of ability, gender, and ethnic backgrounds in the class. The reasons for creating such groups include maximizing the opportunities for peer tutoring and support, improving cross-race and cross-gender relations, ensuring that every group has at least one student who can do the work (Kagan, 1992), and making groups comparable for fair inter-group competition (Slavin, 1990). While much research has found that lower ability students gain more from working in heterogeneous groups than working in homogeneous groups or individually (see Webb & Palincsar, 1996), fewer studies have conducted these comparisons for high-ability students, where researchers typically define high-ability as the top 25% or 33% of a population or sample. Studies comparing individual and group work for these high-ability students yield inconsistent results. In the comparison of homogeneous grouping, heterogeneous grouping, and individual work from which the present study stems, Webb, Nemer, Chizhik, & Sugrue (1998) found that high-ability students performed better in homogeneous groups than in heterogeneous groups, with high-ability students working individually performing in the middle (but not significantly different from students working in either heterogeneous or homogeneous groups). Two other studies, in contrast, found advantages of heterogeneous grouping over individual work for high-ability students for one out of three computer learning posttests (Hooper, Temiyakarn, & Williams, 1993) and on both outcome measures of reading recall and higher level reasoning (Johnson & Johnson, 1993).

A number of other studies compare achievement of high-ability students in heterogeneous and homogeneous groups. In a large majority of these studies, high-

ability students performed equally well on achievement tests after working in heterogeneous and homogeneous groups (e.g., Armstrong, 1994; Azmitia, 1988; Carter & Jones, 1994; Hooper & Hannafin, 1988, 1991; Hooper, Ward, Hannafin, & Clark, 1989; Jones & Carter, 1994; Jones, Carter, & Rua, 1998; Lawrenz, & Munch, 1985; Melser, 1999; Sherman & Klein, 1995; Skon, Johnson, & Johnson, 1981). Most of the remaining studies found higher achievement in homogeneous groups than in heterogeneous groups on some or all outcomes (Baron, 1994; Fuchs, Fuchs, Hamlett, & Karns, 1998; Knupfer, 1993). A notable exception lies in a study that demonstrated that high-ability students learned more in heterogeneous groups than in homogeneous groups (Webb, 1980) because, in heterogeneous groups, high-ability students often assumed the role of teacher and explained the material to other group members (typically the low-ability student). In homogeneous groups, high-ability students exchanged relatively few explanations, apparently because they assumed that everyone could master the material without help.

Some studies reported differences in the behavior of high-ability students in the two group compositions, even when no differences in achievement appeared. Carter and Jones (1994) found that high-ability students in heterogeneous pairs working on activities related to balance interacted in more productive ways than those in homogeneous pairs. The high-ability students in heterogeneous pairs spoke more, took more turns speaking, and exhibited more helping behavior. Similarly, Jones and Carter (1994) found that high-ability students in heterogeneous pairs modeled thinking processes and strategies for manipulating equipment, and guided and monitored low-ability students' progress. Fuchs, Fuchs, Hamlett, and Karns (1998), in contrast, found that high-ability students in heterogeneous pairs did not adopt a teaching role and, for the most part, did the work themselves, possibly because the task proved too complex to involve low-ability students' participation.

The participants in the present study came from a project investigating the effects of group ability composition on group processes and outcomes in science performance assessments (Webb et al., 1998). In that study, students in 21 eighth-grade classes worked on a science assessment in heterogeneous groups, homogeneous groups, or individually, and subsequently worked on another science assessment individually. Below-average students who worked with above-average students showed higher achievement than did below-average students who worked in groups without above-average students. Below-average students learned from the more accurate and higher quality answers and explanations given in heterogeneous

groups. Yet unlike their low-ability peers, high-ability students working in homogeneous groups outperformed high-ability students who worked in heterogeneous groups. The analyses of behavior and experiences of high-ability students in that study, however, failed to uncover explanations for the differences in achievement across group compositions. For example, the quality of the group discussion did not seem to correlate with outcome measures for high-ability students. Furthermore, high-ability students did not engage in more high-level participation (e.g., making suggestions, paraphrasing others' suggestions) in one group composition than in the other, a variable that positively related to learning outcomes in previous research (e.g., Webb et al., 1995). To understand the performance results found in the previous study, the present study focused exclusively on the experiences and outcomes of high-ability students and carried out intensive analyses of high-ability students' interactions with other group members.

Method

Sample

The sample in the original study (Webb et al., 1998) consisted of 662 seventh-grade and eighth-grade students (21 classes) from five schools in Los Angeles County. The present study focuses on the top 25% of that sample, designated as high-ability (described in more detail below). Two schools had small numbers of high-ability students, and no homogeneously high groups, so we did not include them in the analyses. A third school had a small number of homogeneous and heterogeneous groups, but because its overall performance ranked considerably below the two remaining schools, we did not include it in the analyses. The sample for analysis, then, consisted of the high-ability students from the remaining two schools (9 classes; $n = 99$). Of these students, 83 students worked in 47 three-person groups, and 16 students worked individually. The samples from the two schools were predominantly White (89%), with small numbers of Black students (2%), Hispanic students (2%) and Asian students (6%). Females comprised approximately half of the sample (49%).

Procedures

At the beginning of the study, prior to instruction in electricity and electric circuits, all students completed three pretests: vocabulary (Vocabulary Test V-1, Buros, 1974), verbal reasoning (the New Jersey Test of Verbal Reasoning Form B,

Shipman, 1985), and nonverbal reasoning (Raven's Progressive Matrices). After the administration of these tests, all teachers conducted a 3-week unit on electricity and electric circuits in their classrooms. We gave the teachers the freedom to teach the topic in their own way, using their own instructional materials. We told teachers that the assessments administered at the end of instruction would target students' knowledge of the concepts of voltage, resistance, and current, as well as the relationships among them, but teachers did not know in advance what form our evaluation of the students' knowledge of these concepts would take.

At the end of the instructional unit, students completed a hands-on and a written assessment measuring their understandings of voltage, resistance, current, and the interrelationships between these concepts. Students finished both activities individually, without help from other students or from the teacher. The hands-on component required students to assemble manipulatives (batteries, bulbs, wires, and resistors) into electric circuits, draw diagrams of their circuits, and then answer questions (multiple choice and written explanations). Proctors videotaped all assembled circuits. A subset of the paper-and-pencil written component consisted of items analogous to the hands-on element (see Appendix) but required students to draw the circuits instead of assemble them. The other items on the written component did not have direct analogues in the hands-on component, and therefore we did not include them in the present analyses.

One month later, with no intervening instruction or review, students re-completed the same two science assessments (hands-on and written). Students completed the hands-on component first and the written module the next day. For the hands-on component, approximately 80% of the students in each class worked in collaborative three-person groups; as a control, the remaining 20% of the students worked individually at separate desks with no interaction with others. All students completed the written component individually.

When students worked in collaborative groups, each group received one set of manipulatives (batteries, bulbs, wires, resistors), but students had to complete their own written responses (although they could work together on composing the answers). Students received instructions to work together to complete the tasks, to help each other understand how to solve the problems, to ask questions of each other, and to assist group members who seemed confused or appeared to need help. Of the seven or eight groups in each class, proctors videotaped five groups while they completed the test.

Design

The design, then, had three phases. In the first phase, students completed the verbal and nonverbal pretests, and the first administration of the hands-on and paper-and-pencil assessments individually. In the second phase, students completed the hands-on component again either in collaborative three-person groups or individually. In the third phase, all students individually completed the paper-and-pencil component again. Table 1 summarizes the design of the study.

Scoring of Science Achievement Assessments

Hands-on component. The hands-on component consisted of two tasks. For each task, students received a bag of materials containing 1.5-volt and/or 9-volt batteries, wires, bulbs, and graphite resistors (Task 1: two 9-volt batteries and two 1.5-volt batteries, three bulbs in bulb holders, and seven wires with alligator clips; Task 2: two 9-volt batteries, two bulbs in bulb holders, three graphite resistors, and seven wires with alligator clips). Students received instructions to assemble pairs of circuits so that the bulb in one circuit was brighter (or dimmer) than the bulb in the other circuit. After circuit construction, we asked students to draw diagrams of their circuits, explain why one circuit was brighter than the other, and answer three multiple-choice items about which of their two circuits had higher voltage, resistance, and current. Further, the test asked students to write an explanation to justify each multiple-choice answer (see Appendix).

Students assembled a variety of circuits. Because different pairs of circuits gave rise to different correct answers, we scored the multiple-choice items and the explanations according to the circuits that students assembled (as shown on the

Table 1
Design of Study

Phase 1	Phase 2	Phase 3
INDIVIDUAL	GROUP	INDIVIDUAL
Pretests	Hands-on science	Paper-and-pencil science
Verbal reasoning		
Vocabulary		
Nonverbal reasoning		
	OR	
Hands-on science	INDIVIDUAL	
Paper-and-pencil science	Hands-on science	

videotapes of their circuits). For example, if a student assembled two circuits that each contained a nine-volt battery and a 1.5-volt battery, both circuits would have equal voltage. If a student assembled one circuit with one 9-volt battery and two 1.5-volt batteries and the other with only one 9-volt battery, the first circuit would have the higher voltage.

Each multiple-choice item received a score of correct or incorrect (0,1) based on the circuits assembled. Coders scored justifications on a 0-to-1 scale according to accuracy and completeness. For example, when asked "Why was voltage in Circuit A higher than in Circuit B?" coders assigned the following scores:

- 1 if a student mentioned the relative number of batteries in the two circuits in addition to the relative power or voltage generated by the batteries;
- 0.67 if a student mentioned the relative number of batteries in the two circuits but not the strength or voltage of the batteries;
- 0.33 if a student mentioned batteries but not the relative number or relative strength of them; and
- 0 if the explanation was irrelevant or if it displayed confusion over cause and effect (for example, "the voltage is higher because it is brighter").

Two coders scored all justifications. The correlation between coders' ratings exceeded 0.95 for all justification items.

Paper-and-pencil written component. The portion of the paper-and-pencil written component presented in these analyses consists of items exactly analogous to the hands-on component except that students received pictures of the equipment (batteries, bulbs, resistors) instead of actual manipulatives (see Appendix). Otherwise, on this portion of the test the students received identical instructions. We asked students to construct two circuits (draw diagrams using the items given in order to make the bulb in one circuit brighter (or dimmer) than the bulb in the other) and answer the same multiple-choice and explanation questions.

Scores Used in the Analyses

Hands-on component score. The hands-on component score used for analysis consisted of the mean of the three multiple-choice items and four justification items for each of the two tasks. This totaled 14 items (equally weighted). We calculated one hands-on score for Phase 2 (internal consistency $\alpha = .86$) and one for Phase 3 ($\alpha = .83$).

(4) debilitating processes occurring in the group (initiated or received by the high-ability student).

For categories 1 and 2, responsiveness to need for help, we took a two-pronged approach to coding group interactions. We coded the signal that indicated a need for help and coded the response given. Signals included nonverbal behavior (looking confused), general statements of confusion ("I don't know how to do this"), questions that signaled confusion, a lack of comprehension, or a need for information (requests for the answer, an explanation, information, or a justification for why someone gave the answer he or she did), and errors (incorrect multiple-choice answers, incorrect or incomplete justifications, incorrect suggestions for circuit building or diagram drawing). Responses included answers (multiple choice or justification), explanations, information, suggestions, corrections of error, agreement, a non-content response ("I don't know"), and negative responses (insulting, dismissing requests for help). We coded all responses for accuracy, completeness, level of elaboration, and whether the help-giver could have provided a better or more complete response based on his or her knowledge as demonstrated in the Phase 1 measures.

We then placed each combination of signal and response along two continua of responsiveness. The first continuum looked at the content of the response, specifically the accuracy and level of elaboration of the content provided in response to a need for help. The continuum ranged from +3 (fully accurate and elaborated) to -3 (inaccurate and elaborated). Descriptions of each point on the content continuum appear in Table 2. The second continuum took into account the motivation of the student in giving a particular response, that is, the intended helpfulness of the response from the perspective of the help-giver. We coded a response as high on the intended helpfulness continuum if the help-giver gave an elaborated response, whether or not it was correct. This continuum also ranged on a scale from +3 (highly responsive) to -3 (insulting a fellow group member). Descriptions of each point on the intended helpfulness continuum also appear in Table 2.

Category 3 consisted of high-ability students' behaviors when no one in the group overtly signaled a need for help. These behaviors included offering suggestions, answers, explanations, and information; listening to other students make such contributions; agreeing with others' contributions; ignoring others' contributions; and making no contribution when no one else in the group contributed either. We used the frequency of each kind of behavior in the analyses.

Table 2

Continua of Student Responsiveness to Other Students' Need for Help

Level	Description
Content continuum	
+3	Correct answer, complete with explanation
+2	Correct answer that completely answers the question, but with little or no explanation)
+1	Partially correct answer
0	No content
-1	Probably incorrect, very unclear
-2	Incorrect answer with little explanation
-3	Incorrect answer with explanation
Intended helpfulness continuum	
+3	Correct or incorrect response with explanation
+2	Correct or incorrect response with minimal explanation
+1	Correct or incorrect response with no explanation
0	Neutral
-1	Providing a response at a lower level than the capacity of the help-giver
-2	Ignoring or dismissing need for help
-3	Insulting

Finally, category 4 consisted of debilitating behavior: ordering others to do something or to stop doing something, commandeering equipment (including preventing others from participating in circuit construction), giving or receiving insults, interrupting others or being interrupted, and carrying out off-task conversation or behavior. The analyses used the frequencies of these behaviors.

Results and Discussion

Overview of Analysis Strategy

The analyses proceeded in four phases. The first phase examined performance differences across group ability compositions. We looked at group ability composition in three ways. (1) We examined performance differences across high-ability students in the four group compositions defined by the lowest ability in the group (low, low-medium, medium-high, high). (2) We collapsed all heterogeneous groups (groups with high-ability students and at least one student at another ability level—low, low-medium, or medium-high) and compared the performance of high-

Paper-and-pencil analogue score. As described above, we analyzed the subset of the written component analogous to the hands-on component. The only difference between these two modules was that students received pictures of the batteries, bulbs, and resistors instead of the actual equipment. We combined scores on the paper-and-pencil analogue items in the same way as the hands-on component scores to form a composite score (14 items, equally weighted). We calculated one written analogue score for Phase 2 ($\alpha = .85$) and one for Phase 3 ($\alpha = .85$).

Determining Ability Levels

We computed a composite score based on both the pretests and the science tests administered during Phase 1 to serve as a measure of students' competence prior to group-work. In the ability composite, we weighted the science scores more heavily than the verbal and nonverbal reasoning scores because, among students who worked individually in all phases of the study (the total sample), science scores in Phases 2 and 3 had a higher correlation with science scores in Phase 1 (r s ranged from .68 to .69) than with the verbal and nonverbal scores in Phase 1 (r s ranged from .40 to .60). We call the composite measure *ability* to distinguish it from the achievement scores in Phases 2 and 3. Although we use the term ability here, we should note that the composite combines science achievement, verbal and nonverbal reasoning ability, and vocabulary, and so it has a heavy achievement component.

We classified students as low, low-medium, medium-high, or high in ability based on the composite ability score. One quarter of the sample comprised each ability level.

Group Compositions

A variety of group compositions included high-ability students. For the present analyses, we describe the group compositions according to the lowest ability level in the group: low (6 groups), low-medium (10 groups), medium-high (23 groups), and high (8 groups). Proctors videotaped 6 low groups, 9 low-medium groups, 13 medium-high groups, and 8 high groups. Regardless of the ability mix in the group, we formed groups so that they would be heterogeneous on gender and ethnic background to reflect the class mix of student characteristics.

We used stratified random sampling from the different ability levels in order to select the 20% of students in each class who worked individually during the second phase. In the two schools analyzed here, 16 high-ability students worked individually during that phase.

Coding of Videotapes

Group discussions. Proctors videotaped groups as the students worked collaboratively on the hands-on component. Coders then assigned two group discussion scores using the same scoring scheme applied to the hands-on test: the accuracy of students' multiple-choice answers and the quality of their justifications. For each multiple-choice item on the test, coders gave the group a 1 if anyone in the group voiced the correct answer and a 0 otherwise. For each item, coders gave the group a score for the highest level of justification articulated based on the contributions of the three students in the group. For example, when looking at a justification about why one circuit had higher current than the other, if one student explained that Circuit A had higher current because it had more voltage and another student explained that lower resistance also led to higher current, coders combined the two justifications (one about voltage and one about resistance) to form a single justification that coders then scored. The justification score, then, reflected the maximum level of discussion in the group and could have reflected the contributions of one, two, or all three students.

To check interrater agreement, two coders scored the group discussion variables for 42 groups (2 per classroom). The raters' codes correlated at 1.00 for accuracy of answers and 0.99 for quality of justifications.

It should be noted that the coding of these group discussion variables differed somewhat from the coding in prior analyses (Webb et al., 1998). In the earlier study, if a group did not provide any verbal response for an item, we coded the accuracy of the response 0. In the present study, if a group did not provide a verbal response for an item, we did not code the accuracy; the code would be missing for that item. The accuracy of group discussion codes, then, reflects only the accuracy of the responses given verbally.

Individual behavior of students in groups. In order to evaluate how working in collaborative groups impacted students, we looked at four categories of student behavior: (1) the responsiveness of the high-ability student to other students' need for help (as indicated by questions, errors, or nonverbal signals of confusion), (2) the responsiveness of the other members of the group to the high-ability student's need for help, (3) the contributions made by the high-ability student in the absence of any indication of a need for help on the part of other group members, and